WATER RESOURCES RESEARCH, VOL. 37, NO. 11, PAGES 2749-2759, NOVEMBER 2001

Toward improved streamflow forecasts: Value of semidistributed modeling

Douglas P. Boyle,[1] Hoshin V. Gupta, and Soroosh Sorooshian
Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

Victor Koren, Ziya Zhang, and Michael Smith
Hydrology Laboratory, Office of Hydrologic Development, NOAA/NWS, Silver Spring, Maryland, USA

Abstract. The focus of this study is to assess the performance improvements of semidistributed applications of the U.S. National Weather Service Sacramento Soil Moisture Accounting model on a watershed using radar-based remotely sensed precipitation data. Specifically, performance comparisons are made within an automated multicriteria calibration framework to evaluate the benefit of "spatial distribution" of the model input (precipitation), structural components (soil moisture and streamflow routing computations), and surface characteristics (parameters). A comparison of these results is made with those obtained through manual calibration. Results indicate that for the study watershed, there are performance improvements associated with semidistributed model applications when the watershed is partitioned into three subwatersheds; however, no additional benefit is gained from increasing the number of subwatersheds from three to eight. Improvements in model performance are demonstrably related to the spatial distribution of the model input and streamflow routing. Surprisingly, there is no improvement associated with the distribution of the surface characteristics (model parameters).

1. Introduction, Motivation, and Scope

Conceptual rainfall-runoff (CRR) models are often difficult to calibrate because of the large number of functional parameters and complex relationships. Systematic manual calibration techniques such as those developed by the U.S. National Weather Service (NWS) for calibration of the Sacramento Soil Moisture Accounting (SAC-SMA) model can result in good model calibrations but are complicated and highly labor intensive. Traditional automatic calibration procedures take advantage of the speed and power of digital computers, while being objective and relatively easy to implement. However, they do not provide parameter estimates and hydrograph simulations that are considered acceptable by the hydrologists responsible for operational forecasting and have therefore not entered into widespread use. *Gupta et al.* [1998] acknowledged some of the more serious limitations with the classical automatic approach and presented a more general multicriteria framework for model calibration that recognizes the multiobjective nature of the problem. *Yapo et al.* [1998] developed the MOCOM-UA algorithm, an effective and efficient methodology for solving the multiobjective global optimization problem. *Boyle et al.* [2000] analyzed the similarities and differences between the automatic and manual calibration approaches and proposed a new hybrid multicriteria approach that combines the strengths of each. They demonstrated that the new approach could be used to emulate some of the important aspects of the manual approach pursued by an expert. Further, the approach provides information that is useful for evaluating the limitations of the various structural components of the model, thereby pointing toward potential structural improvements.

The majority of CRR model calibration studies have typically been concerned with "lumped" applications (i.e., averaging of the dominant subwatershed-scale processes that contribute to the overall watershed-scale response). However, remotely sensed, high-resolution hydrologic information (e.g., Next Generation Weather Radar (NEXRAD) Stage III precipitation data [*Klazura and Imy*, 1993]) is now becoming widely available (at least in Europe and the United States), and hydrologists have begun to incorporate this information into their modeling procedures [e.g., *Smith et al.*, 1999; *Michaud and Sorooshian*, 1994; *Ogden and Julien*, 1993, 1994; *Winchell et al.*, 1998; *Krajewski et al.*, 1991]. This has led to the development of relatively complex "fully distributed" models that allow the user to construct a very detailed representation of the spatial variability of the hydrologic processes within the watershed (e.g., Sŷsteme Hydrologique European (SHE) [*Abbott et al.*, 1986], among others). The main premise is that the spatial detail will lead to an improved understanding of the watershed behavior while improving simulations. However, the degree to which the spatial variability of each process needs to be represented (to provide the maximum improvement in the simulations) is not well understood. Further, these models typically have a very large number of parameters for which values must be estimated, either through a calibration procedure or from maps of watershed properties.

In the case of streamflow (flood) forecasting the primary concern is to generate estimates of discharge at a limited number of fixed points along the river channel, and there is typically little need for spatially detailed information about the various

[1]Now at Desert Research Institute, University and Community College System of Nevada, Reno, Nevada, USA.

Paper number 2000WR000207.
0043-1397/01/2000WR000207$09.00

internal states of the watershed. An attractive alternative to the lumped and fully distributed approaches is the so-called semidistributed approach in which the watershed is conceptualized as a network of functionally distinguishable land segments (subwatersheds), each represented by a lumped CRR model. The runoff response from each segment is routed to the outlet of the watershed to compute the total watershed response. In this approach the spatial variability of the hydrologic processes within the watershed is represented through the number and location of segments, which determine the degree of spatial distribution of the model input (precipitation), structural components (soil moisture and streamflow routing computations), and surface characteristics (parameters).

The added complexity in going from a lumped to a semidistributed representation of a watershed results in a significantly more complex model calibration problem. CRR models tend to use complex, highly parameterized, vertical representations (of the movement of moisture through the soil), and the overall number of parameters to be estimated for each unit can be quite large. Further, many of the parameters may not be supported (identifiable) by the information contained within the observed data, remotely sensed or otherwise [*Jakeman and Hornberger*, 1993; *Wagener et al.*, 1999; *Wheater et al.*, 1993]. As the number of hydrologic segments is increased, the calibration procedure (manual or automatic) can quickly become intractable. Operational use of semidistributed rainfall-runoff model applications to produce flood forecasts has therefore been limited.

The focus of this paper is to provide an assessment of the potential improvements in rainfall-runoff model performance that can be achieved by semidistributed modeling of a watershed using radar-based (NEXRAD) remotely sensed precipitation data. The relative benefits of spatially distributing the model input (precipitation), structural components (soil moisture and streamflow routing), and surface characteristics (parameters) are examined. The CRR model used is the NWS Sacramento Soil Moisture Accounting (SAC-SMA) model [*Burnash et al.*, 1973]. The multicriteria framework developed by *Boyle et al.* [2000] for application to lumped hydrologic models is used to calibrate the semidistributed model in terms of three objective measures designed to reflect the different observable characteristics of watershed behavior (peak flow and timing, quick recession, and base flow). Multicriteria performance comparisons among the different model applications are used to evaluate the benefits of various types and degrees of spatial complexity. Results from an independent NWS manual (expert) calibration study are used as a basis to evaluate the approach.

This paper is organized as follows: the background and context for the work is discussed in section 2. The theoretical and practical basis for applying the multicriteria methodology to investigate spatially distributed modeling approach on a study watershed is presented in section 3. The results of the model applications are presented in section 4, and the results and future extensions of the study results are discussed in section 5.

2. Background

Since the early 1990s, the NWS has been investigating methods to assess the benefit (improvement in streamflow forecasts) of representing the spatial variability of the high-resolution precipitation, soil, and vegetation properties within the NWS modeling system [*Smith et al.*, 1999, 2000]. One primary focus has been the incorporation of high-resolution, remotely sensed, Next Generation Weather Radar Stage III precipitation data into standard NWS procedures for flood forecasting using the SAC-SMA model.

The SAC-SMA model is a continuous soil moisture accounting algorithm with an upper zone representing the upper soil layer and interception storage and a lower zone representing the majority of the soil moisture and the longer groundwater storage. Within each zone, water is stored in two forms (tension and free water). The model computes six components of flow: direct runoff, surface runoff, interflow, supplementary base flow, primary base flow, and subsurface outflow. The details of these computations have been discussed previously in the literature [e.g., *Burnash et al.*, 1973; *Burnash*, 1995; *Peck*, 1976; *Brazil and Hudlow*, 1981; *Sorooshian and Gupta*, 1983]. Each component of flow, except subsurface outflow (considered a loss from catchment), contributes directly to the channel inflow and may be routed to the outlet of the catchment using a unit hydrograph.

The soil moisture accounting component of the model has 17 parameters whose values must be specified (Table 1). While some of these parameters can be related to observable characteristics of the watershed, many are abstract conceptual representations of nonmeasurable watershed characteristics that are difficult to estimate and are therefore typically specified through a calibration procedure. In addition, the unit hydrograph ordinates must be derived either from observed precipitation and streamflow information or from empirical methods related to physical characteristics of the watershed.

In a recent study the NWS applied the SAC-SMA model to five different watersheds in both lumped and semidistributed applications and compared the resulting streamflow simulations to observed data [*Smith et al.*, 1999, 2000; *Zhang et al.*, 2001]. The results indicated that for four of the watersheds characterized as having a deep soil layer (>150 cm) the accuracy of the semidistributed simulations "did not show significant improvement in accuracy" when compared to the lumped simulations. However, the results for the remaining watershed (Blue River at Blue, Oklahoma), characterized as having much shallower soils, "were improved significantly" by use of the semidistributed model. On the basis of these findings the Blue River watershed was selected for the studies reported in this paper.

2.1. Description of Blue River Watershed and Data

The Blue River watershed is located in southern Oklahoma near Blue, Oklahoma (Figure 1). The watershed is a long narrow valley with a contributing area of 1227 km^2 distributed primarily along the main channel. The soils are generally characterized as shallow (<2 m) sandy clay throughout the region. The mean annual precipitation is 1003 mm, and the runoff coefficient is estimated (from annual data) at 0.20 [*Niadas*, 1999]. The U.S. Geological Survey (USGS) surface water discharge station 07332500, Blue River near Blue, Oklahoma, is located at the outlet of the watershed and has been in operation since 1948.

A NEXRAD Stage III precipitation data set for the Blue River watershed was obtained for the period June 1, 1993, to April 30, 1999. The NEXRAD product offers gridded precipitation estimates, spatially averaged over 4 km by 4 km grid cells and temporally averaged over 1 hour. For the same time period, hourly estimates of instantaneous discharge are available from the USGS surface water discharge station. The

Table 1. Parameters and State Variables of the SAC-SMA Model

Parameter	Description	Multicriteria Calibration Range	NWS Lumped
UZTWM	upper zone tension water maximum storage (mm)	1.0–200.0	40
UZFWM	upper zone free water maximum storage (mm)	1.0–100.0	37
LZTWM	lower zone tension water maximum storage (mm)	90–200.0	160
LZFPM	lower zone free water primary maximum storage (mm)	50–400.0	120
LZFSM	lower zone free water supplemental maximum storage (mm)	1.0–90.0	75
ADIMP	additional impervious area (decimal fraction)	0.0–0.4	0.0
UZK	upper zone free water lateral depletion rate (day^{-1})	0.1–0.9	0.5
LZPK	lower zone primary free water depletion rate (day^{-1})	0.0008–0.01	0.002
LZSK	lower zone supplemental free water depletion rate (day^{-1})	0.01–0.2	0.03
ZPERC	maximum percolation rate (dimensionless)	1.0–250.0	180
REXP	exponent of the percolation equation (dimensionless)	1.0–3.0	1.8
PCTIM	impervious fraction of the watershed area (decimal fraction)	0.0–0.01	0.008
PFREE	fraction of water percolating from upper zone directly to lower zone free water storage (decimal fraction)	0.0–0.6	0.04
RIVA	riparian vegetation area (decimal fraction)	0.00	0.00
SIDE	ratio of deep recharge to channel base flow (dimensionless)	0.00	0.00
RSERV	fraction of lower zone free water not transferable to lower zone tension water (decimal fraction)	0.30	0.30
PXMLT	precipitation multiplication factor (dimensionless)	1.00	1.00
UZTWC	upper zone tension water storage content (mm)		
UZFWC	upper zone free water storage content (mm)		
LZTWC	lower zone tension water storage content (mm)		
LZFPC	lower zone free primary water storage content (mm)		
LZFSC	lower zone free secondary water storage content (mm)		
ADIMC	additional impervious area content (mm)		

NWS-derived hourly values of potential evapotranspiration (PET) are based on long-term average values obtained from the atlas of National Oceanic and Atmospheric Administration (NOAA) free water evaporation [*NOAA*, 1982]. It should be noted that the PET values vary seasonally but not annually (i.e., an estimate of PET for 1 year has been made and is repeated for each year of record).

The NWS created a digital elevation model (DEM) of the Blue River watershed from 100 m (cell size) elevation data. The watershed was partitioned into eight subwatersheds (Figure 1) on the basis of an analysis of DEM stream connectivity data (stream channel structure), and the variability of the high-resolution soil property information available from the U.S. Department of Agriculture (USDA) State Soil Geographic Database (STATSGO) for the resulting subwatersheds (V. Koren, Hydrologic Laboratory, NWS, personal communication, 2000). The average soil depth determined from the STATSGO information and the contributing area of each subwatershed are listed in Table 2.

Mean areal precipitation values for each of the eight subwatersheds were estimated from the 4 × 4 km NEXRAD Stage III hourly precipitation data. Unit hydrographs for each subwatershed were developed in conjunction with the DEM, using the methodology described by *Smith et al.* [1999], to route the simulated channel inflow to the outlet of the watershed. For the lumped conceptualization the unit hydrograph was derived from the subwatershed unit hydrographs.

2.2. NWS Modeling Strategy

The NWS applied the SAC-SMA model in both lumped and semidistributed (eight subwatersheds) forms to the Blue River watershed. In the lumped case, the channel inflow was computed at each time step for the entire watershed and then routed to the outlet with a single unit hydrograph. In the semidistributed case the soil moisture computations were made separately for each subwatershed, and the resulting simulated channel inflows were then routed independently to the outlet of the watershed and combined to compute the total simulated streamflow for the entire watershed.

The NWS first used a sophisticated, highly interactive manual procedure to estimate values for 13 of the SAC-SMA parameters (four were set to default values) [*Anderson*, 1997] for the lumped watershed case. In this procedure, initial values for the calibration parameters are estimated from analytical relationships derived by *Koren et al.* [2000] based on soil property data (STATSGO soil texture data). While these estimated values are not "optimum" values, they are considered (by the NWS) to be "very reasonable initial approximations" [*Koren et al.*, 2000]. Next, a systematic sequence of steps is followed to develop parameter estimates based on an examination of the hydrological database (precipitation, PET, and streamflow) of the watershed. Periods in the observed time series data are identified where specific hydrologic processes are dominant (e.g., base flow, interflow, surface flow, evaporation, transpiration, abstraction, infiltration, etc.). For each of these periods the closeness of the model and the data is evaluated visually, and values for the relevant parameters are estimated by heuristic methods. In the last (and most difficult) step of the procedure the hydrologist must deal with the effects of parameter interactions on the model responses by simultaneously evaluating a number of subjective and objective criteria while iteratively adjusting the parameter values so that the model matches the behavior of the watershed system as closely as possible.

For the semidistributed case, there are 104 parameters to be estimated (13 for each of the eight subwatersheds). For each subwatershed, initial parameter values were estimated from the soils data using the analytical relationships developed by *Koren et al.* [2000]. Next, the calibration parameter values were adjusted manually while maintaining the ratios between the initial parameter values among the subwatersheds. Finally, a

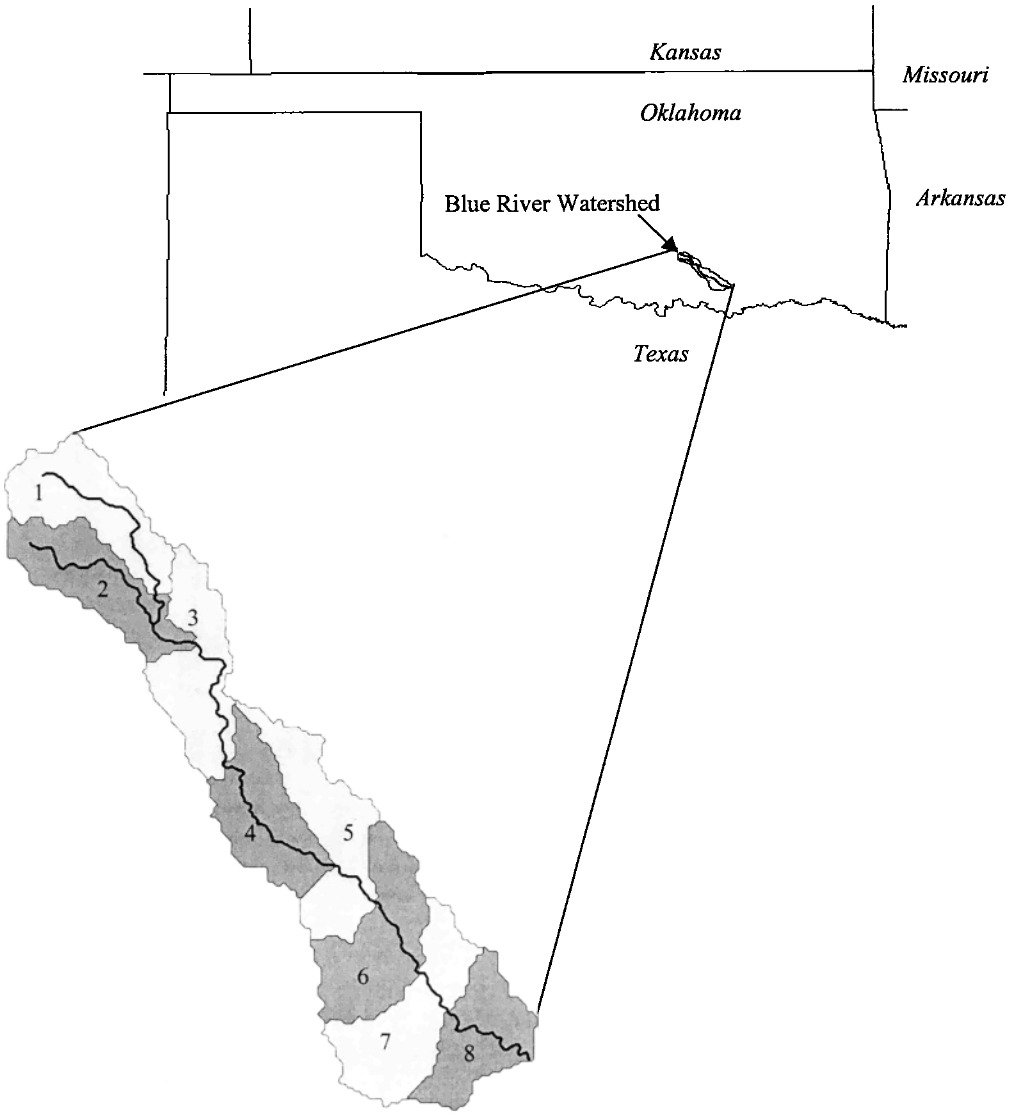

Figure 1. Location of Blue River watershed.

"fine-tuning" of the parameter estimates for each subwatershed was done through a complicated and subjective examination of the hydrological database of the watershed and through consideration of the soil moisture state and simulated hydrograph (at the outlet of the watershed) for each subwatershed. This step requires the hydrologist to search the entire simulation period for consistent deviations between the model and watershed behavior and to make appropriate parameter adjustments (e.g., precipitation events that occur primarily over a single subwatershed can be isolated and the corresponding parameters adjusted to improve the associated streamflow simulations). Clearly, the large number of model parameters and limited number of isolated events for each subwatershed limit the effectiveness of this approach.

Table 2. Subwatershed Attributes

Subwatershed	Contributing Area, km^2	Average Soil Depth,m
1	153.1	1.19
2	150.0	0.88
3	153.1	0.83
4	144.0	1.70
5	162.9	1.71
6	165.0	1.75
7	169.9	1.75
8	129.0	1.75

3. Methods

The manual calibration conducted by the NWS was used in this study as the basis for an evaluation of the strengths and weaknesses of the automatic multicriteria calibration approach [*Gupta et al.*, 1998; *Boyle et al.*, 2000]. In addition, a number of different model runs were conducted to investigate the benefit of different levels of spatial representation of model input (precipitation), structural components (soil moisture and streamflow routing computations), and surface characteristics (parameters) of the SAC-SMA model applied to the Blue River watershed. The study was designed to complement the NWS semidistributed studies on the Blue River by expanding our understanding of the specific benefits associated with different levels of spatial representation of the model.

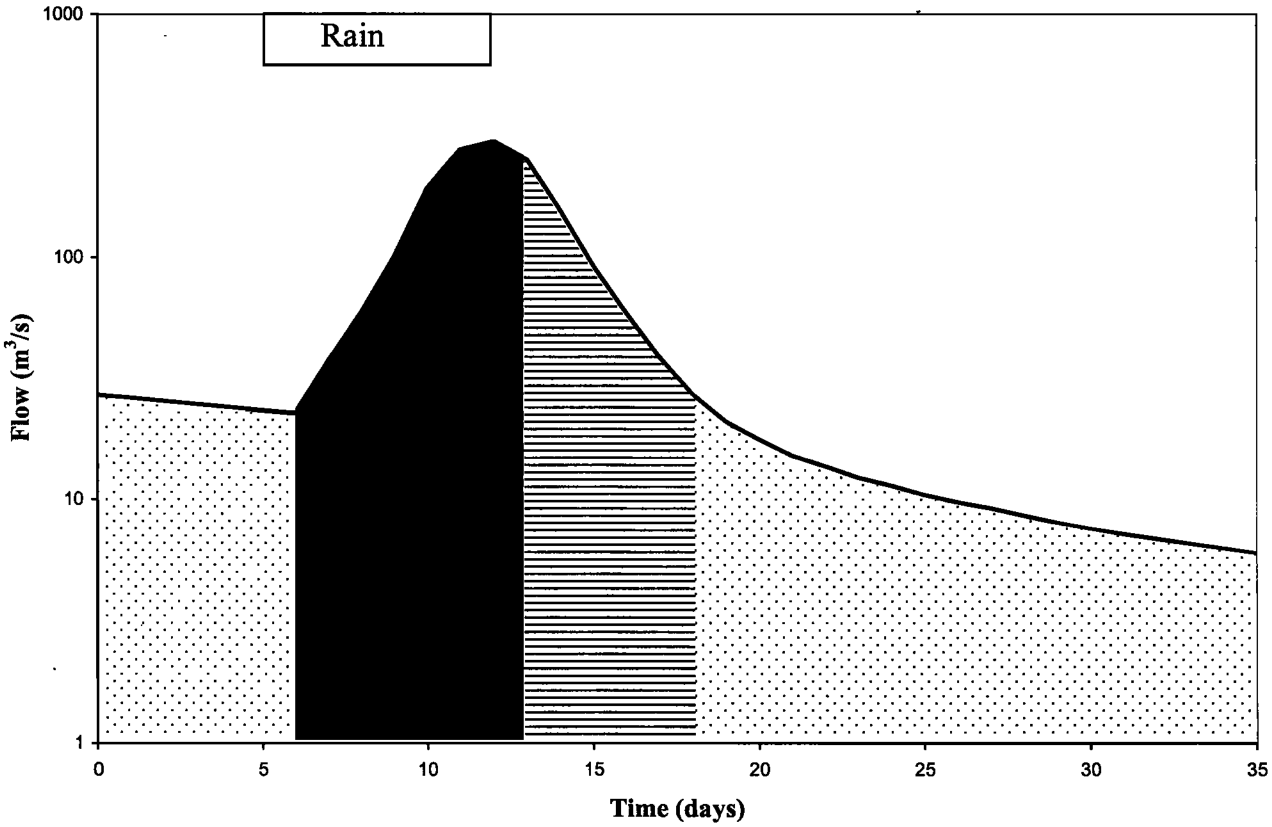

Figure 2. Partitioning of the observed hydrograph into three components: driven Q_D (solid black area), nondriven quick Q_Q (horizontally lined area), and nondriven slow Q_S (dotted area).

3.1. Multicriteria Parameter Estimation Methodology

The multicriteria approach to calibration presented in detail by *Boyle et al.* [2000] combines the strengths of both automated and manual calibration methods. The approach involves the identification of several characteristic features of the observed streamflow hydrograph, each representing a distinct (preferably unique) aspect of the behavior of the watershed. In brief, the hydrograph is partitioned into three components based on the reasonable assumption that the behavior of the watershed is inherently different during periods "driven" by rainfall and periods without rain. Further, the periods immediately following the cessation of rainfall and dominated by interflow can be distinguished from the later periods that are dominated by base flow. The streamflow hydrograph can therefore be partitioned into three components (Figure 2), which we call "driven" (Q_D), "nondriven quick" (Q_Q), and "nondriven slow" (Q_S). The time steps corresponding to each of these components are identified through an analysis of the precipitation data and the time of concentration for the watershed. The time steps with nonzero rainfalls, lagged by the time of concentration for the watershed, are classified as driven. Of the remaining (nondriven) time steps those with streamflows lower than a certain threshold value (e.g., mean of the logarithms of the flows) are classified as nondriven slow, and the rest are classified as nondriven quick. For each of the components the closeness between the model outputs and the corresponding observed values is estimated separately using the RMSE statistic, resulting in three evaluation criteria, designated as FD (driven), FQ (nondriven quick), and FS (nondriven slow), respectively.

An important characteristic of the multiobjective problem is that it does not, in general, have a unique solution. Because of errors in the model structure (and other possible sources), it is not usually possible to find a single unique solution that simultaneously minimizes all of the criteria. Instead, it is common to have a "Pareto set" of solutions with the property that moving from one solution to another results in the improvement of one criterion while causing a deterioration in one or more others. The Pareto set represents the minimum uncertainty that can be achieved for the parameters via calibration, without subjectively assigning relative weights to the individual model responses. The size and properties of this set are related to errors in the model structure and data. In this study, the Multi-Objective Complex evolution algorithm [*Yapo et al.*, 1998; *Bastidas et al.*, 1999] was used to solve the multicriteria optimization problem. MOCOM is a general-purpose multiobjective global optimization algorithm that provides, in a single optimization run, a set of points that approximate the Pareto set. For details, the reader is referred to *Gupta et al.* [1998] and *Yapo et al.* [1997, 1998].

3.2. Description of This Study

In this study, a series of lumped and semidistributed applications of the SAC-SMA model to the Blue River watershed was made to investigate the improvements in model performance associated with various levels of spatial representation of model input (precipitation), structural components (soil moisture and streamflow routing computations), and surface characteristics (parameters). Each model application was designed to isolate the effects of the different levels of spatial

Table 3. Table of Different Modeling Cases[a]

Case	Case Name	Precipitation P	Soil Moisture Computation S	Routing Computation R	Parameters θ
1	LUMP-ALL	L	L	L	L
2	DIST-PS	D	D	L	L
3	DIST-PSR	D	D	D	L
4	DIST-SR	L	D	D	L
5	DIST-PSRθ	D	D	D	D

[a]L, lumped; D, distributed.

representation in terms of specific desirable watershed behaviors (driven flow, "peaks and timing"; nondriven quick flow, "quick recession" responses; and nondriven slow, "base flow" responses). The calibration data set (precipitation, PET, and streamflow) used in this study was the same as that used in the NWS study. Model calibration and evaluation of the performance improvements for each application were performed using the multicriteria approach described above. For each case the Pareto optimal solution space for the three criteria (FD, FQ, and FS) was estimated by 500 solutions generated using the MOCOM algorithm.

Five separate cases of spatial distribution of the SAC-SMA model were investigated (see Table 3). In case 1 (LUMP-ALL) the SAC-SMA model was applied in a lumped configuration (precipitation P, soil moisture computations S, and streamflow routing computations R, were all lumped) to the Blue River watershed. This case served as a benchmark for performance comparisons with cases 2–5, in which the SAC-SMA model was applied in varying levels of spatial distribution to the eight-subwatershed configuration used by the NWS. Note that in cases 2–4 the model parameters were treated as lumped (all the subwatersheds were assigned the same values of the 13 calibration parameters) and only the spatial distribution of the model input and structural components was investigated.

In case 2 (DIST-PS) the precipitation and soil moisture computations were spatially distributed among the subwatersheds, but the routing was treated as lumped. In this application, soil moisture computations were performed separately to compute separate channel inflow sequences for each subwatershed, but these were combined into a total channel inflow for the entire watershed before routing to the outlet of the watershed using a single unit hydrograph. In case 3 (DIST-PSR) the precipitation, soil moisture computations, and streamflow routing computations were spatially distributed among the subwatersheds to assess the additional benefit of distributed routing. In this application, the channel inflow computed for each subwatershed was independently routed to the outlet of the watershed with separate unit hydrographs and then combined to estimate the total runoff from the watershed. In case 4 (DIST-SR) the precipitation was treated as lumped over the entire watershed, but the soil moisture and streamflow routing computations were spatially distributed among the subwatersheds. This configuration was designed to investigate the value of the spatially distributed precipitation through comparison with cases 1–3. In case 5 (DIST-PSRθ) the additional value of spatially distributing the watershed soil properties was investigated by allowing some of the model parameters to vary among the subwatersheds. The five parameters allowed to vary were LZTWM, REXP, UZTWM, UZFWM, and UZK, selected on the basis of the results of previous studies by the NWS [*Koren et al.*, 2000] indicating that these parameters exhibit empirical relationships with different soil property data.

Finally, to further investigate the effects of spatial representation, cases 2–5 were repeated using a smaller number of subwatersheds (i.e., the entire watershed was partitioned into a three-subwatershed configuration). In this new configuration (also provided to us by the NWS), the original subwatersheds 1, 2, and part of 3 were combined to form the new subwatershed 1 of the three-subwatershed configuration. Similarly, 4, 5, and parts of 3 and 6 were combined to form the new subwatershed 2, while 7, 8, and part of 6 were combined to form the new subwatershed 3. The mean areal precipitation and PET for each of the three new subwatersheds were estimated by the NWS using the same methods mentioned previously.

4. Results

4.1. NWS Manual Calibration Results

The NWS manual calibration studies were used as benchmarks for evaluation of the automatic calibration studies described above. The manual calibration results are shown in the multicriteria format in Figures 3a–3c. Figures 3a–3c present the results for each case using two-dimensional projections of the three-criteria solution space (NWS lumped case is large open square and NWS semidistributed case is large open circle). Clearly, the semidistributed application results in an improvement in the model's ability to simulate the observed flow in terms of FQ and FS, as compared with the lumped application. There is a slight decrease, however, in the model's ability to simulate the driven flows measured by FD.

4.2. Automatic Multicriteria Calibration Results

4.2.1. Case 1: Lumped model. The results of the multicriteria automatic calibration of the lumped case (LUMP-ALL) are also shown in Figures 3a–3c, as a three-criteria trade-off surface represented by the set of 500 Pareto optimal solutions (indicated by the light shaded dots). The inability of the model to simultaneously match all three aspects of the hydrograph is clearly illustrated. For example, Figure 3b illustrates the smoothly varying trade-off between the model's ability to match the driven (Q_D) and the nondriven slow (Q_S) portions of the hydrograph (similarly, see Figure 3c and, to a lesser extent, Figure 3a).

A visual comparison of the 500 Pareto solutions with the NWS lumped solution (open square) in Figures 3a–3c shows that the automatic approach provides a closer fit of the base flow responses (FS) and, to a lesser extent, the quick recession responses (FQ). In terms of the peaks and timing (FD), however, most of the 500 Pareto solutions are inferior to the NWS lumped solution.

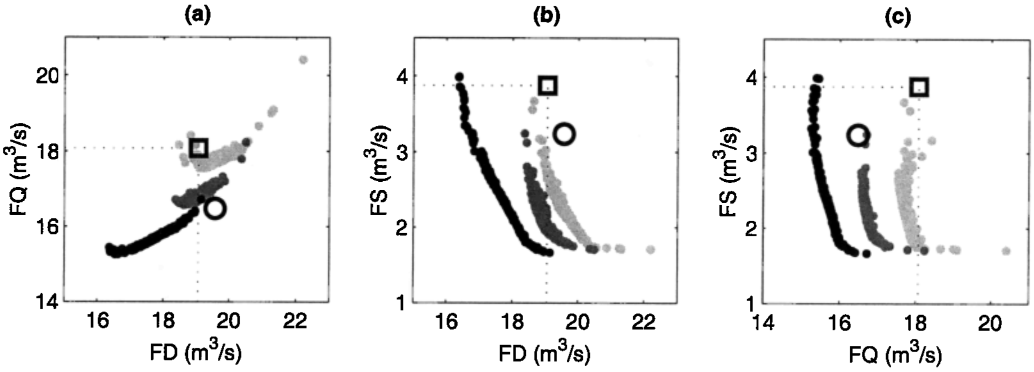

Figure 3. Pareto solutions obtained with the automatic multicriteria approach to calibrate the Sacramento Soil Moisture Accounting (SAC-SMA) model: (a–c) two-dimensional projections of objective space. Dots correspond to 500 Pareto solutions for case 1 (light shaded dots), case 2 (dark shaded dots), and case 3 (solid dots). U.S. National Weather Service (NWS) manual calibration results are shown for lumped (open square) and semidistributed (open circle).

The variability in the parameter values across the 500 Pareto optimal solutions for the case LUMP-ALL is shown in Figure 4a (shaded lines) using a normalized parameter plot. Each line across the graph represents one of the parameter sets. The maximum range for each parameter represents the range over which the multicriteria calibration procedure was performed (see Table 1). Notice that the multicriteria optimization has resulted in a significant reduction in the parameter range. Notice also that the seven parameters LZFPM, LZFSM, LZPK, LZSK, PCTIM, ZPERC, and REXP show only small amounts of variability while corresponding well with the NWS lumped solution (solid line) obtained by an expert via manual calibration. The relative lack of variability in these parameters indicates that they are not primarily responsible for the performance trade-offs associated with the inability of the model to simultaneously match the three different components of the hydrograph (indicated by FD, FQ, and FS). Of these parameters, the first five (LZFPM, LZFSM, LZPK, LZSK, and PCTIM) are considered to be the ones that are "most identifiable" and relatively easy to estimate via careful off-line examination of the observed hydrograph and precipitation data [*Peck*, 1976]. The fact that the automatic multicriteria approach gives com-

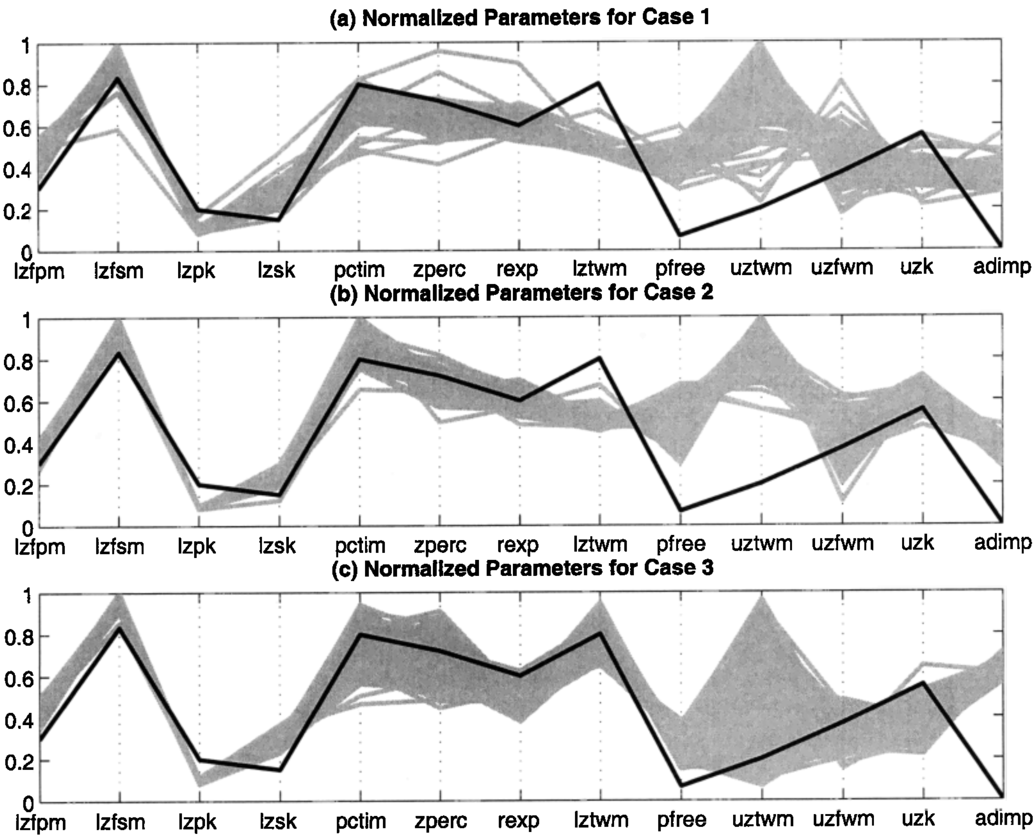

Figure 4. Normalized parameters for Pareto solutions obtained with the automatic multicriteria approach to calibrate the SAC-SMA model: (a) case 1, (b) case 2, and (c) case 3.

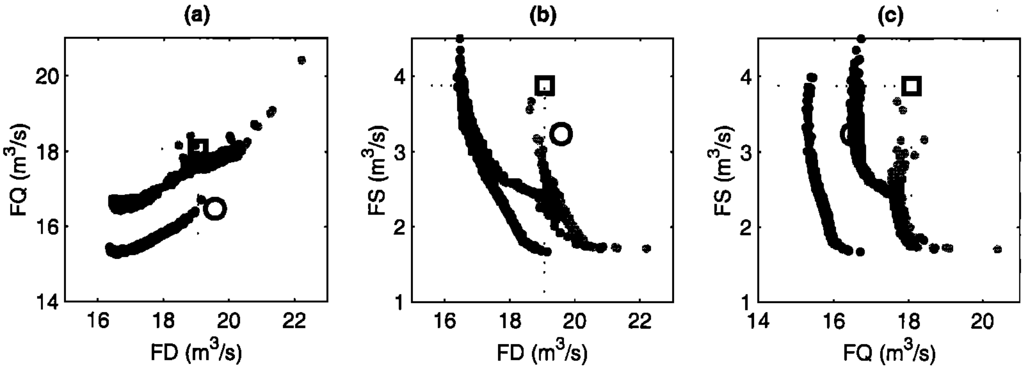

Figure 5. Pareto solutions obtained with the automatic multicriteria approach to calibrate the SAC-SMA model: (a–c) two-dimensional projections of objective space. Dots correspond to 500 Pareto solutions for case 1 (light shaded dots), case 4 (dark shaded dots), and case 3 (solid dots). NWS manual calibration results are shown for lumped (square) and semidistributed (open circle).

parable values for these parameters is a good indication of the reasonableness of the automatic approach. Note, however, that the two parameters, ZPERC and REXP, that control the behavior of the percolation (infiltration) component of the model are not easy to estimate directly from the observed data.

There is, in general, larger variability in the estimates for the remaining six model parameters (and, in particular, UZTWM and UZFWM), which are primarily related to the properties of the upper soil layer and control the partitioning of the hydrograph into different quick flow components (overland and interflow). This suggests that the model performance tradeoffs indicated in Figures 3a–3c are associated primarily with inadequacies in the representation of the near-surface soil processes.

4.2.2. Case 2: Distributed input and soil moisture. The results of the multicriteria automatic calibration of case 2 (DIST-PS) are shown in Figures 3a–3c. Note that, in this case, the channel inflows for all the subwatersheds are lumped together and routed to the outlet using a single unit hydrograph. The results for the eight-subwatershed configuration did not give better results than the three-subwatershed configuration. Therefore the results presented here will draw primarily from the results of the three-subwatershed study. Comparison of the solutions for this case (dark shaded dots) with the lumped case (case 1, LUMP-ALL, light shaded dots) indicates a significant benefit to allowing the precipitation input and the soil moisture computations to be distributed. In particular, the ability of the model to simulate the quick recession responses (FQ) and, to a lesser extent, the peaks/timing (FD) has been improved. However, there appears to be no additional impact on the model's ability to simulate the base flow responses (FS).

A visual comparison of the DIST-PS results with the NWS lumped solution in Figures 3a–3c clearly shows that the automatic approach provides a closer fit to the observed data in terms of all three criteria FD, FQ, and FS. Further, comparison of the DIST-PS results with the NWS semidistributed solution shows that most of the 500 Pareto solutions provide a better fit to the base flow (FS) and peaks/timing (FD), while providing a comparable fit to the quick recession (FQ).

The variability in the parameter values across the 500 Pareto optimal solutions for the DIST-PS case is shown in Figure 4b. Notice that the variability in the parameters has generally decreased when compared with case 1 (LUMP-ALL). As before, the range for the seven parameters, LZFPM, LZFSM, LZPK, LZSK, PCTIM, ZPERC, and REXP, is close to the values for the lumped NWS solution, further supporting the notion that the automatic multicriteria approach is finding appropriate values for these parameters.

4.2.3. Case 3: Distributed input, soil moisture, and routing. The results of the multicriteria automatic calibration of case 3 (DIST-PSR) are also shown in Figures 3a–3c. In this case the precipitation, soil moisture computations, and channel routing are all treated separately for each subwatershed. Again, the results for the eight-subwatershed configuration did not give better results than the three-subwatershed configuration, and results are therefore only presented for the latter configuration. The 500 Pareto optimal parameter sets (solid dots) show that routing the channel inflow independently from each subwatershed to the outlet of the watershed improves the model's ability to simulate both the quick recession responses (FQ) and the peaks/timing (FD). Once again, there is no additional improvement in the model's ability to simulate the base flow responses (FS). A visual comparison of the 500 Pareto solutions for this case with the NWS lumped and semidistributed solutions (Figures 3a–3c) clearly shows that the automatically calibrated semidistributed model DIST-PSR provides a much better reproduction of the watershed response in terms of all three criteria, FD, FQ, and FS.

The normalized parameter values for the 500 Pareto solutions are shown in Figure 4c. Again, the variability in the parameters is generally less than in case 1 (LUMP-ALL), and the parameters LZFPM, LZFSM, LZPK, LZSK, PCTIM, ZPERC, and REXP are close to the lumped NWS solution. However, the variability of parameter UZTWM (upper zone tension water maximum capacity) significantly increased when compared with the results from cases 1 and 2, suggesting that this parameter may need to take different values in each subwatershed.

4.2.4. Case 4: Distributed soil moisture and routing: Lumped input. The results of the multicriteria automatic calibration of case 4 (DIST-SR) are shown in Figures 5a–5c. In this case the input was treated as lumped (averaged over the watershed), thereby simulating the lack of availability of radar-based precipitation data, while the model computations were run in semidistributed mode. Only the three-subwatershed case is shown. The 500 Pareto optimal parameter sets obtained for this case (dark shaded dots) are shown overlaid on the results from case 1 (LUMP-ALL, light shaded dots) and case

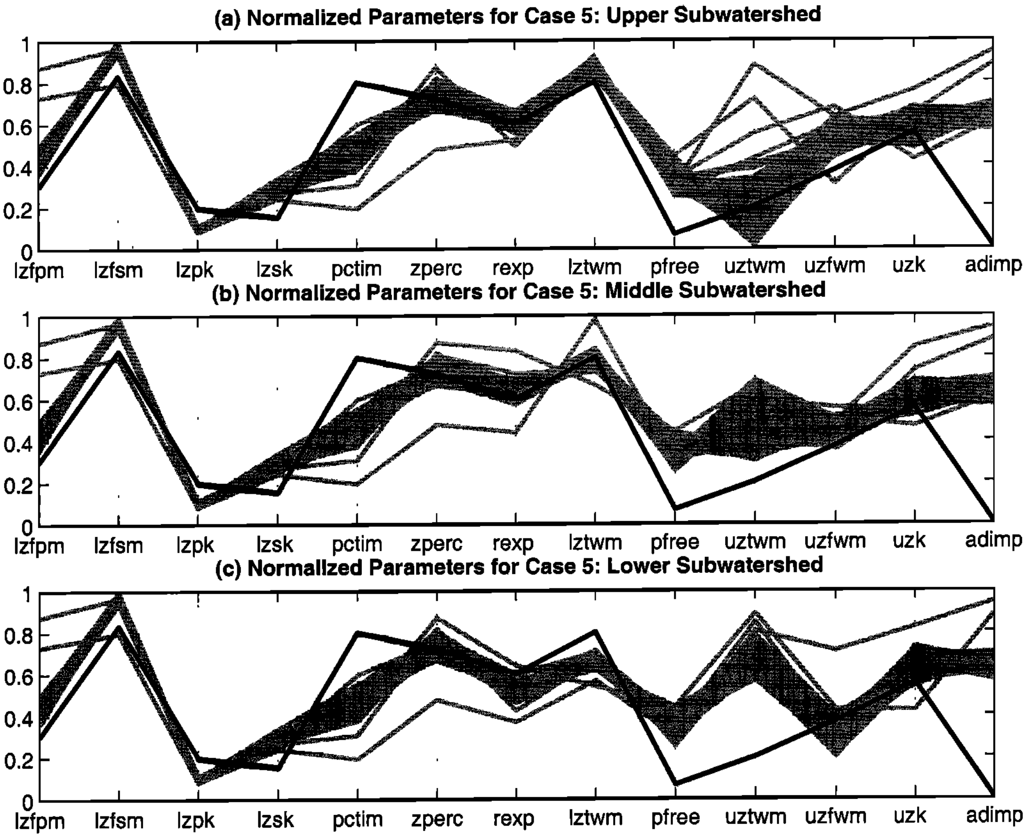

Figure 6. Normalized parameters for Pareto solutions obtained with the automatic multicriteria approach to calibrate the SAC-SMA model for case 5: (a) upper subwatershed, (b) middle subwatershed, and (c) lower subwatershed.

3 (DIST-PSR, solid dots). Note that FQ for most of the case 4 solutions is improved over case 1, but not as "good" as case 3, indicating that spatially distributed routing and input (precipitation) both contribute to improvements in simulating the quick recession (FQ). However, the case 4 solutions (lumped input and distributed routing) that provide the best match to FD (peaks/timing) are comparable to the case 3 solutions (distributed input and distributed routing), indicating that the simulation of the peaks is controlled more strongly by the spatial representation of routing than by the spatial representation of the inputs (precipitation).

4.2.5. Case 5: Distributed input, soil moisture, routing, and parameters. In case 5 (DIST-PSRθ) the additional value of spatially distributing the watershed soil properties was investigated by allowing five of the model parameters (LZTWM, REXP, UZTWM, UZFWM, and UZK) to vary among the subwatersheds. Unexpectedly, the calibration results showed no noticeable criterion value improvements over the solutions obtained for the lumped parameter case 3 (DIST-PSR). This is surprising, given the larger number of calibration parameters (23) compared with the number of calibration parameters (13) in case 3 (DIST-PSR). Additional runs made with other combinations of parameters also provided no noticeable gains. However, the parameter plots shown in Figures 6a–6c (each subplot represents a different subwatershed) indicate that the calibration has converged to significantly different values for the parameter UZTWM (upper zone tension water maximum capacity) in each subwatershed: low value for the upper subwatershed, medium value for the middle subwatershed, and high value for the lower subwatershed. This variability is consistent with results obtained under case 3, where forcing this parameter to be lumped results in a large trade-off range (i.e., the parameter must take different values to fit different portions of the data record). Further, it seems reasonable that the increasing trend for UZTWM from the upper to lower sections of the watershed is hydrologically realistic and is consistent with soil depth maps for the watershed.

4.3. Comparison of Automatic and Manual Hydrographs

Examples of the simulated hydrographs for a 120 hour portion of the calibration period are shown in Figure 7. The solid dots correspond to the observed data, the light shaded region corresponds to the hydrograph trade-off uncertainty associated with the case 3 (DIST-PSR) multicriteria solution, and the dashed and solid lines correspond to the NWS lumped and semidistributed results, respectively. Because the ordinates of the unit hydrographs were kept fixed (not dynamically adjusted) during the calibration, the hydrograph recessions are not well simulated. Nevertheless, the multicriteria solutions are able to reproduce the timing and magnitude of the flood peak extremely well.

5. Summary and Conclusions

The semidistributed approach to modeling the spatial variability of important hydrologic processes is becoming popular as the availability of high-resolution hydrologic information continues to increase. However, the spatial detail with which

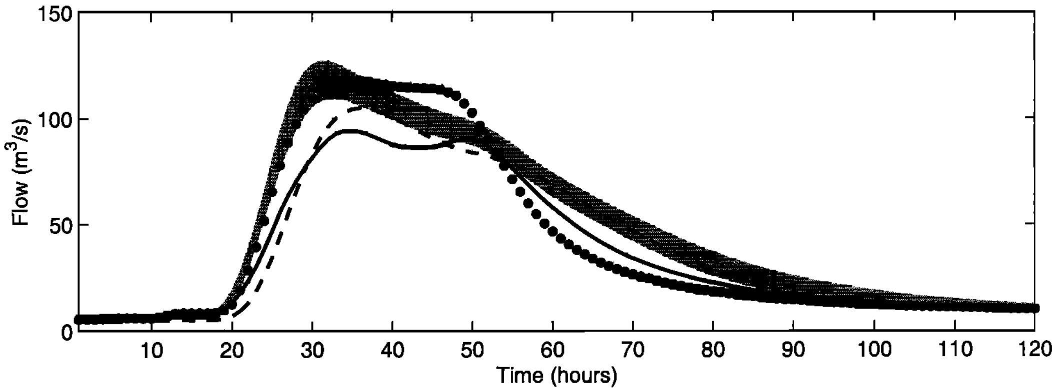

Figure 7. Hydrograph range associated with the Pareto solution set for case 3. The dots correspond to the observed streamflow data, the light shaded region corresponds to the 500 Pareto solutions, the dashed line corresponds to the NWS lumped solution, and the solid line corresponds to the NWS semidistributed solution.

this variability needs to be represented to provide accurate streamflow simulations is not well understood. Further, the increased complexity of the semidistributed approach (and, in particular, the large number of parameters) results in a calibration problem of considerable difficulty. Traditional manual and automatic calibration approaches typically provide a single suboptimal solution, with little or no information about the uncertainty in the estimated parameters or the model performance.

The applicability of multicriteria automatic calibration methods to the calibration of semidistributed watershed models has been demonstrated in this study. The automatic approach was shown to be able to effectively handle the increased model complexity and provide parameter estimates (and hence model performance) that are comparable or superior to those obtained by the manual/expert approach. The improved results are also achieved at a considerable savings in "calibration time": after problem setup, the automatic approach requires ~2–4 hours of computing time on a Sun workstation, compared with several days to weeks of person time for the manual approach (M. Smith, Hydrology Laboratory, NWS, personal communication, 2000). In addition, the study has demonstrated how multicriteria methods provide a useful framework for the systematic investigation of appropriate model complexity.

The study was specifically motivated by the needs of the U.S. National Weather Service for improved procedures for calibration of watershed models in the context of the Advanced Hydrologic Prediction System modernization initiative. In particular, it was designed to complement ongoing NWS studies into the development of semidistributed modeling strategies that exploit the increasing availability of spatial hydrologic information, including NEXRAD precipitation. The effectiveness and efficiency of the automatic approach enabled us to rapidly investigate the specific benefits associated with different levels of spatial representation of various model components, including the model input (precipitation), structural components (soil moisture and streamflow routing computations), and surface characteristics (parameters). In particular, it was found that (1) the semidistributed model provided significant performance improvements over the lumped model, (2) there was a limit to the performance improvements associated with increasing representation of spatial hydrologic variability in the model, (3) the main improvements were provided by spatial representation of the precipitation (inputs) and structural components (soil moisture and streamflow routing computations), (4) little or no improvements were provided by spatial representation of the soil properties (model parameters), (5) the spatial variability in hydrologic information contributed mainly to improved simulation of the flood peaks and the quick recessions, and (6) semidistributed modeling provided no improvements in representation of the base flow. These results are, of course, specific to the SAC-SMA model for the Blue River watershed, although conclusions 1, 2, and 6 may prove to be more generally applicable.

Conclusion 4, relating to the lack of improvement in model performance when the model parameters are allowed to be distributed spatially, is both unexpected and interesting. This seems contrary to the general belief among hydrologists that the spatial variability in soil properties exerts a significant control on the hydrologic response of a watershed. For the Blue River watershed at least, our results indicate that the dominant control on the hydrologic response is the watershed topography (which strongly determines the routing characteristics and (perhaps less strongly) the spatial distribution of precipitation) and that the impacts of variations in the soil and vegetation properties are averaged out by the time the streamflow reaches the watershed outlet. Further studies to investigate the conditions under which the various components of hydrologic variability do (or do not) contribute to variations in total watershed response would provide significant improvements to our understanding of watershed behavior.

Research aimed at further understanding the relationship between model complexity and performance improvement is ongoing. This includes investigation of the benefits associated with various levels of vertical model complexity. While the scope of this study was limited to calibrations on a single watershed, further testing on a large number of watersheds is planned through participation in the Distributed Model Intercomparison Project recently proposed by the Hydrology Laboratory of the NWS. The results of that work will be reported in due course. As always, we invite dialog with others interested in these topics.

Acknowledgments. Partial financial support for this research was provided by the National Science Foundation (EAR-9418147), the Hydrologic Research Laboratory of the National Weather Service (grants NA47WG0408 and NA77WH0425), the Sustainability of Water Resources in Semi-Arid Regions (SAHRA), and by the National Aeronautics and Space Administration (NASA-EOS grant NAGW2425).

References

Abbott, M., J. Bathrust, J. Cunge, P. O'Connell, and J. Rasmussen, An introduction to the European Hydrological System-Systeme Hydrologique European (SHE), 1, History and philosophy of a physically-based, distributed modelling system, *J. Hydrol.*, *87*, 45–59, 1986.

Anderson, E. A., Hydrologic model calibration using the Interactive Calibration Program (ICP), Hydrol. Res. Lab., U.S. Natl. Weather Serv., Silver Spring, Md., 1997.

Bastidas, L. A., H. V. Gupta, and S. Sorooshian, The Multi-Objective COMplex evolution algorithm, MOCOM-UA: User's guide, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson, 1999.

Boyle, D. P., H. V. Gupta, and S. Sorooshian, Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674, 2000.

Brazil, L. E., and M. D. Hudlow, Calibration procedures used with the National Weather Service Forecast System, in Water and Related Land Resource Systems: *IFAC Symposium, Cleveland, Ohio, USA, 28–31 May 1980*, edited by Y. Y. Haimes and J. Kindler, pp. 457–466, Pergamon, New York, 1981.

Burnash, R. J. C., The NWS river forecast system: catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo., 1995.

Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, A generalized streamflow simulation system: Conceptual modeling for digital computers, U.S. Dep. of Commer., Natl. Weather Serv. and State of Calif., Dep. of Water Resour., Sacramento, Calif., 1973.

Gupta, H. V., S. Sorooshian, and P. O. Yapo, Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763, 1998.

Jakeman, A. J., and G. M. Hornberger, How much complexity is warranted in a rainfall-runoff model?, *Water Resour. Res.*, *29*(8), 2637–2649, 1993.

Klazura, G. E., and D. A. Imy, A description of the initial set of analysis products available from the NEXRAD WSR-88D system, *Bull. Am. Meteorol. Soc.*, *74*(7), 1293–1311, 1993.

Koren, V. I., M. Smith, D. Wang, and Z. Zhang, Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, paper presented at 15th Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif., 2000.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S. C. Jain, A Monte Carlo study of rainfall sampling effect on a distributed catchment model, *Water Resour. Res.*, *27*(1), 119–128, 1991.

Michaud, J. D., and S. Sorooshian, Effect of rainfall-sampling errors on simulations of desert flash floods, *Water Resour. Res.*, *30*(10), 2765–2775, 1994.

National Oceanic and Atmospheric Administration (NOAA), Atlas of NOAA free water evaporation, *Tech. Rep. NWS 33*, Silver Spring, Md., 1982.

Niadas, I. A., Rainfall-runoff model regionalisation: An application to Midwest USA, M.S. thesis, Dep. of Civ. and Environ. Eng., Imp. Coll., London, England, 1999.

Ogden, F. L., and P. Y. Julien, Runoff sensitivity to temporal and spatial rainfall variability at runoff plane and small basin scales, *Water Resour. Res.*, *29*(8), 2589–2597, 1993.

Ogden, F. L., and P. Y. Julien, Runoff model sensitivity to radar rainfall resolution, *J. Hydrol.*, *158*, 1–18, 1994.

Peck, E. L., Catchment modeling and initial parameter estimation for the national weather service river forecast system, *NOAA Tech. Memo. NWS HYDRO-31*, Off. of Hydrol., Natl. Weather Serv., Silver Spring, Md., 1976.

Smith, M. B., V. Koren, D. Johnson, B. D. Finnerty, and D. J. Seo, Distributed modeling: Phase 1 results, *NOAA Tech. Rep. NWS 44*, 210 pp., Natl. Hydrol. Res. Lab., Silver Spring, Md., 1999.

Smith, M. B., V. Koren, Z. Zhang, D. Wang, and S. Reed, Semi-distributed versus lumped model simulations: Comparisons using observed data for RFC scale basins, *Eos Trans. AGU*, *81*(19), Spring Meet. Suppl., Abstract H32F-03, 2000.

Sorooshian, S., and V. K. Gupta, Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, *Water Resour. Res.*, *19*(1), 260–268, 1983.

Wagener, T., M. J. Lees, and H. S. Wheater, A generic rainfall-runoff modeling toolbox, *Eos Trans., AGU*, *80*(17), Fall Meet. Suppl., F203, 1999.

Wheater, H. S., A. J. Jakeman, and K. J. Beven, Progress and directions in rainfall-runoff modeling, in *Modelling Change in Environmental Systems*, edited by A. J. Jakeman, M. B. Beck, and M. J. McAleer, pp. 101–132, John Wiley, New York, 1993.

Winchell, M., H. V. Gupta, and S. Sorooshian, On the simulation of infiltration- and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation, *Water Resour. Res.*, *34*(10), 2655–2670, 1998.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, A multiobjective global optimization algorithm with application to calibration of hydrologic models, *HWR Tech. Rep. 97-050*, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson, 1997.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, Multi-objective global optimization for hydrologic models, *J. Hydrol.*, *204*, 83–97, 1998.

Zhang, Z., V. Koren, M. Smith, S. Reed, and D. Wang, Comparison of continuous lumped and semi-distributed hydrologic modeling using NEXRAD data, *J. Hydrol. Eng.*, in press, 2001.

D. P. Boyle, H. V. Gupta, and S. Sorooshian, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA. (dboyle@dri.edu; hoshin@hwr.arizona.edu)

V. Koren, M. Smith, and Z. Zhang, Hydrology Laboratory, Office of Hydrologic Development, NOAA/NWS, Silver Spring, MD 20910, USA. (victor.koren@noaa.gov)

(Received January 15, 2001; revised May 5, 2001;
accepted May 24, 2001.)